Contribution ID: **32**   Type: **Oral**

Exposing and facilitating the searching of fusion data using an Open API-based data acquisition system supporting multiple data-access methods

Tuesday 6 July 2021 16:00 (10 minutes)

Fusion-related experiments (WEST, MAST, JET, etc...) produce large amounts of data. In the future, we can expect ITER to produce an equally large amount of raw data coming from each and every shot.

Even though (fusion community as a whole) collects large sets of data, scientists suffer from a lack of an amalgamated shot catalogue and different access requirements. At the moment, finding data for a given shot requires having access to all sites (where experiments are run), being able to use different data formats (HDF5, MDSPlus, raw data), and having access to storage locations where databases are kept. It is also not possible to search for given physical characteristics of data as there is no single, unified format of storage or ontology.

Catalogue QT 2 and the Fair4Fusion Dashboard aims to solve some of these issues. By utilizing the IMAS data format and storing a reduced description of experiments' results (metainformation is stored inside the so-called Summary IDS), we provide scientists with a convenient way of browsing, searching, and (in the future) obtaining experimental data. By combining information from various sources (MAST, WEST, etc...) and different acquisition techniques, UDA (Universal Data Access), MDSPlus files, text based file formats, etc..., we are able to provide users with a consistent view and search functionality of data coming from different sources. By developing and combining loosely coupled components, based on Web Services, we can present data to users not only via a dedicated user interface (Web Application created with ReactJS) but also via command line tools and Jupyter Notebook based scripts. Openly available APIs allow connections to third party applications as long as they belong to the same Federated Authentication and Authorisation Infrastructure. Thanks to moving authorization and authentication responsibilities to external Identity Providers, we were able to move user management out of the scope of the application itself. This way, we can provide multiple, independent installations of Catalogue QT 2 and benefit from the common Authentication and Authorisation Infrastructure infrastructure. Installation of Catalogue QT 2 is possible in virtually any environment - due to the fact that we provide both bare metal-based installations and Docker-based components.

This paper presents the architecture of a proposed solution, ways of combining different, loosely-coupled components coming from different projects and possible future directions of Catalogue QT 2's evolution.

Keywords: IMAS, MDSPlus, data acquisition, fusion experiments, data analysis, web services, Docker, Authentication and Authorisation Infrastructure

Speaker's Affiliation

Poznan Supercomputing and Networking Center, IBCh PAS, Poznan

Member State or IGO

European Union

Primary author: Mr OWSIAK, Michal (Poznan Supercomputing and Networking Center, IBCh PAS)

Co-authors: Mr PLOCIENNIK, Marcin (Poznan Supercomputing and Networking Center, IBCh PAS); Mr PALAK, Bartek (Poznan Supercomputing and Networking Center, IBCh PAS); Ms FILIPCZAK, Agata (Poznan Supercomputing and Networking Center, IBCh PAS); Mr GRABOWSKI, Piotr (Poznan Supercomputing and Networking Center, IBCh PAS); Mr POGODZINSKI , Bartlomiej (Poznan Supercomputing and Networking Center, IBCh PAS); Mr GALGANSKI, Aleksander (oznan Supercomputing and Networking Center, IBCh PAS); Mr ZOK, Tomasz (Poznan Supercomputing and Networking Center, IBCh PAS); Mr CUMMINGS, Nathan (United Kingdom

Atomic Energy Authority); Mr GIBBONS, George (United Kingdom Atomic Energy Authority); Mr NUNN, Timothy (United Kingdom Atomic Energy Authority)

Presenter: Mr OWSIAK, Michal (Poznan Supercomputing and Networking Center, IBCh PAS)

Session Classification: Database techniques for Information

Track Classification: Database Techniques for Information Storage and Retrieval